2017

# Forging on-campus connections to enhance undergraduate student reasoning, writing, and research skills

Belinda Archibong

Harrison Dekker
*University of Rhode Island,* hdekker@uri.edu

***See next page for additional authors***



Citation/Publisher Attribution
Belinda Archibong, Harrison Dekker, Nathan D. Grawe, Martha L. Olney, Carol Rutz & David Weiman (2017) Forging on-campus connections to enhance undergraduate student reasoning, writing, and research skills, The Journal of Economic Education, 48:4, 317-326, DOI: 10.1080/00220485.2017.1353466
Available at: https://doi.org/10.1080/00220485.2017.1353466

# FORGING ON-CAMPUS CONNECTIONS TO ENHANCE UNDERGRADUATE STUDENT REASONING, WRITING, AND RESEARCH SKILLS

Belinda Archibong, Assistant Professor of Economics, Barnard College, 212-854-8952, Barnard Hall 239, Department of Economics, Barnard College, 3009 Broadway, New York, NY 10027, ba2207@columbia.edu

Harrison Dekker, Associate Professor and Data Services Librarian, University of Rhode Island, 401-874-5935, University Libraries, University of Rhode Island, 15 Lippitt Road, Kingston RI 02881, hdekker@uri.edu

Nathan D. Grawe, Professor of Economics and Former Director, Quantitative Inquiry, Reasoning, and Knowledge (QuIRK) Initiative, Carleton College, (507) 222-5239, Willis Memorial Hall 304, Carleton College, One North College Street, Northfield, MN 55057, ngrawe@carleton.edu

(Corresponding Author) Martha L. Olney, Adjunct Professor of Economics, University of California, Berkeley, 510-642-6083, 530 Evans Hall #3880, Department of Economics, University of California, Berkeley, CA 94720-3880, olney@berkeley.edu

Carol Rutz, Director of the College Writing Program and Senior Lecturer in English, Carleton College, 507-222-4082, Weitz Center for Creativity 124, Carleton College, One North College Street, Northfield, MN 55057, crutz@carleton.edu

David Weiman, Alena Wels Hirschorn '58 Professor of Economics and Faculty Director of the Empirical Reasoning Center, Barnard College, 212-854-5775, Department of Economics, Barnard College, 3009 Broadway, New York, NY 10027, dweiman@barnard.edu

**Corresponding Author**: Martha Olney

# FORGING ON-CAMPUS CONNECTIONS TO ENHANCE UNDERGRADUATE STUDENT REASONING, WRITING, AND RESEARCH SKILLS

Belinda Archibong, Barnard College

Harrison Dekker, University of Rhode Island

Nathan D. Grawe, Carleton College

Martha L. Olney, University of California, Berkeley

Carol Rutz, Carleton College

David Weiman, Barnard College

May 2017

Short Running Title: Forging On-Campus Connections

Abstract:

Research and writing are a critical component of an undergraduate education. Partnerships between economics faculty and campus resources can improve student research and writing skills. Here we describe programs at three different campuses that bridge department and campus resources – the Empirical Reasoning Lab at Barnard College, the Writing Program at Carleton College, and the Library's Data Lab at U.C. Berkeley. We describe each program's mission and structure, provide examples of its impact on student learning, and discuss administrative factors (and hurdles) to consider in implementing similar programs elsewhere.

(88 words)

Key words: Undergraduate Research, Teaching, General Economics, Library Science, Quantitative Reasoning, Writing Center

JEL Classification Codes:
**A220** Economic Education and Teaching of Economics: Undergraduate

# FORGING ON-CAMPUS CONNECTIONS TO ENHANCE UNDERGRADUATE STUDENT REASONING, WRITING, AND RESEARCH SKILLS


## ABSTRACT

Research and writing are a critical component of an undergraduate education. Partnerships between economics faculty and campus resources can improve student research and writing skills. Here we describe programs at three different campuses that bridge department and campus resources – the Empirical Reasoning Lab at Barnard College, the Writing Program at Carleton College, and the Library's Data Lab at U.C. Berkeley. We describe each program's mission and structure, provide examples of its impact on student learning, and discuss administrative factors (and hurdles) to consider in implementing similar programs elsewhere.


Empirical data skills, writing, and research are key components of an undergraduate education in economics. Faculty possess expertise in their areas of specialty. But is that expertise enough when the goal is to teach empirical reasoning? Or writing? Or locating, downloading, cleaning, and merging datasets for research? It should be apparent to economists well-versed in the concept of comparative advantage that student learning is enhanced when faculty work with other campus professionals whose expertise complements our own. In this article, we offer three examples of on-campus units that partner with academic departments to support student learning.

At Barnard College, a liberal arts college for women located in New York City, the Empirical Reasoning Lab supports all departments in including an "empirical reasoning" component in their undergraduate curricula. David Weiman and Belinda Archibong of Barnard's Economics Department note that the Empirical Reasoning Lab started as a division within the library, was moved administratively to become a center due to strong demand for its services, and will soon be a part of Barnard's new Teaching and Learning Center. Students take advantage of individual consultation appointments and faculty can receive assistance in devising curriculum that develops and strengthens empirical reasoning skills.

Nathan Grawe (Economics) and Carol Rutz (Director of the College Writing Program) describe the College Writing Program at Carleton College, a liberal arts college in Northfield, Minnesota. Grawe and Rutz (2009) argue for integrating an institution's quantitative reasoning requirement with its college writing program by framing quantitative reasoning as a form of argument. At Carleton College, quantitative reasoning is integrated with writing across the curriculum in the Quantitative Inquiry, Reasoning, and Knowledge (QuIRK) Initiative. The College Writing Program works with faculty to design effective writing assignments that teach quantitative reasoning. In this contribution, Grawe and Rutz describe particular assignments for Economics principles courses that integrate writing and QR.

The Library's Data Lab at the University of California, Berkeley supplements the usual library undergraduate research support described in Hensley, Shreeves, and Davis-Kahl (2014). The Data Lab offers a physical space for students to work and employs a data librarian, someone with a background in programming and social science research. The data librarian assists library patrons with all aspects of obtaining and preparing to use data and other digital materials. In their contribution, Martha Olney (Economics) and Harrison Dekker (The Library) discuss the work of a data librarian, give examples of student projects made possible by the data librarian's expertise and assistance, and provide guidance on steps to take to bring a data librarian to your campus.

## THE EMPIRICAL REASONING CENTER AT BARNARD COLLEGE

Launched in January 2013, Barnard's Empirical Reasoning Center (ERC) is an essential support of a broader curricular initiative to embed "empirical reasoning" throughout Barnard's undergraduate curriculum, notably in introductory and general education courses. In the Barnard context, "empirical reasoning" means "thinking critically with data" both qualitative and

quantitative.[1] At the most basic level, empirical reasoning refers to the capacity to comprehend and systematically evaluate the varied kinds of empirical arguments and evidence that students routinely encounter in reading assignments, in the popular press, and on websites. It also encompasses "research design" -- the successive steps in developing and carrying out an independent research project including the formulation of an empirical question, specification of the appropriate method of analysis, collection (and even production) of the relevant data, and interpretation and presentation of the results.

Barnard's ERC provides students and faculty from across the College with a variety of services. The lab operates a drop-in help desk staffed by six undergraduate fellows and two graduate assistants, and an online tutorial web page (so far covering Excel, Stata, SPSS, and Qualtrics). The Lab staff and the Associate Director also offer individual student consultations by appointment, totaling over 400 visits (just over 15% of all students) in the 2015-16 academic year. The vast majority of students' queries focused on software applications, but they also raised more general questions about appropriate strategies for visualizing and statistically analyzing their data.

Additionally, the Center devotes significant resources, particularly the efforts of the Associate Director, to collaborate with faculty in developing curricular content for their classes and in offering instructional workshops. Workshops are most often taught in conjunction with curricular innovations, but have also been tailored to faculty's research interests and the training of their undergraduate and graduate research staff. In the 2015-16 academic year including Barnard's Summer Research Institute (SRI), the Center offered 30 workshops with attendance totaling over 900 (or about 1/3 of all) students. In a recent introductory economics class, for example, the Center worked with the instructor to develop a simple empirical exercise analyzing

changes in the crude oil and refined gasoline markets from 1950 to 2013. Through dedicated workshops, the Center's staff instructed students in using spreadsheet tools to manage their data and to hone their data visualization skills. The staff also offered small group sessions for students eager to learn how to derive and interpret basic descriptive statistics (e.g., correlation coefficients, standard deviations etc.). For our interdisciplinary Economic and Social History program, the ERC staff provided critical help desk support for one of our signature courses, Measuring History, in which students with diverse academic backgrounds learn how to apply economic-demographic analyses (such as life tables and Lorenz curves) to illuminate empirically major historical debates. For the SRI, by contrast, the Center offered more advanced courses in Excel, ArcGIS, Matlab, Python and SPSS.

Since its inception the ERC has been jointly funded by a five-year grant from the Mellon Foundation and by the College. At full staffing, its operating budget for the 2015-16 academic year just exceeded $200,000, not including the College's in-kind contribution such as course release time for the Faculty Director. The bulk of its budget covers the wages and salaries of the Associate Director and student staff.

Originally conceived as a division within the library, the "lab" was initially managed by a part-time data (and personal) librarian. In response to soaring student and faculty demands for its services and in anticipation of its move to a larger, dedicated space in Barnard's new teaching and learning center (under construction[2]), the "lab" was elevated to a center under the leadership of a Faculty Director who reports to the Provost and a full-time Associate Director. The Center's leadership continues to collaborate closely with library staff, but also regularly consults with a loosely configured "advisory" committee consisting of faculty and staff who rely on its services. These ties were particularly vital in the search for and recruiting a new Associate Director during

the summer of 2015 and subsequently in expanding the Center's programmatic reach, most recently in diffusing open source statistical and GIS software into specialized methods and upper-level elective courses in economics and other quantitative disciplines.

Two cases presented below illustrate the value of institutional infrastructure such as the ERC in diffusing empirical reasoning throughout the undergraduate curriculum. The first case was designed for a general education economic history course, but can be easily adapted to comparable classes in other empirical subjects. Through a structured assignment, it illustrates the value of simple exploratory methods as a guide to further inquiry, either theoretical or empirical. It also serves as a useful diagnostic tool testing students' proficiency in data analysis and spreadsheet programs, which they tend to overstate in our surveys administered to workshop participants.

The second case is based on an undergraduate Economics senior thesis on the impact of temperature changes on migration in Nepal in which the student combined traditional econometric methods with spatial statistics, GIS and other mapping techniques learned through the ERC. In this instance the Center enabled a student to launch a new avenue of independent empirical research and to develop capacities and interests beyond the scope of the typical undergraduate curriculum.

## Example 1: Rudimentary Demographic Analysis in Economic History

The Economic History assignment asks students to prepare a brief research proposal based on their analysis of nineteenth-century mortality trends in five U.S. cities. Students are given an unbalanced panel of annual death rates, and then guided through the successive steps of this rudimentary exploratory investigation. They must first explain how the rates for each city are

measured and identify any limitations or possible biases in the underlying sources. They then attend an ERC Excel workshop on managing and displaying the data and on the derivation and interpretation of simple linear trend lines. More than mere technical instruction, the workshop and follow-on help-desk sessions show students how to refine their analyses to derive meaningful results. Given the wide variation in mortality experiences across cities and over time, for example, they quickly recognize the shortcomings of graphing all data on a single chart or in drawing linear trend lines over the entire period.

Consistent with the course's learning goals, the students apply their historical knowledge to remedy these problems employing a more inductive approach with Excel's graphical and statistical tools. Many decide to limit their study to a subset of cities or a narrower time period, and are expected to provide compelling historical reasons to justify their choices. In the process they learn yet another valuable set of research-design skills and are able to better empirically analyze the problem. For example, one student charted crude death rates in Boston and New York from 1811 to 1870. She chose these Northeastern cities despite differences in size because they were both major commercial centers and ports of entry for the mass migrations of Germans and Irish beginning in the 1830s. Based on the evidence she posed a number of questions for further research: 1) Were vibrant port cities more vulnerable to cholera epidemics as evidenced by the frequent spikes in mortality rates especially in New York?; 2) What explains the widening gap between mortality rates in the two cities after 1825?; and 3) Why did the mortality rates in the two cities rise above their trend levels in the late pre-Civil War period?

**Example 2: Examining The Impacts of Climate Change on Migration in Nepal**

Providing individual guidance in research design is one of the strengths of the Empirical

Reasoning Center. One example of the Center's effectiveness is the experience of an Economics undergraduate student who, for her senior thesis, wanted to examine the impacts of climate change on migration in Nepal. First, the help desk staff worked with the student to narrow her project down to 3 specific goals:

1. Visualizing the data by mapping the displacement of population due to migration in Nepali states along with population distribution for multiple years
2. Mapping the temperature profile of Nepal over multiple years
3. Determining if there was a statistically significant relationship between change in temperature and migration in Nepal.

Next, the ERC staff assisted the student by identifying and teaching her the relevant statistical tools and methods.

First, the ERC staff taught her the relevant tools in ArcGIS and alternative freely available programs like GeoDa that allowed her to visualize her population and temperature data in a single map. They then taught her exploratory techniques including ones that leveraged the visual display of her data, such as spatial autocorrelation statistics like the global Moran's I to complement the related but non-spatial Pearson correlation coefficient which she learned in her statistics and econometrics classes (Anselin 1995). Finally, ERC staff worked with the student to ensure that the traditional and more novel spatial econometrics models she employed, such as the Spatial Error Model, were sound and that she understood the logic and interpretation of these models (Pace and LeSage 2009).

As a result of these individual consultations provided through the ERC help desk, the student successfully completed her thesis project. With a newfound interest in the intersection of Economics and Regional Science, she later applied and was admitted into a Geography graduate

program.

### Summary

As these two cases illustrate, the Barnard Center allows for a more decentralized pedagogic approach to embedding empirical reasoning in the undergraduate curriculum at the introductory and advanced levels. In lieu of required introductory courses, this model relies on inserting carefully designed empirical modules into popular general education classes, which can expose students to the varieties of empirical evidence and of rudimentary data analysis and so prepare them a standard methods class. To lower the barriers of adoption to what are admittedly more challenging and labor-intensive projects, Center staff assist faculty in developing and testing meaningful but feasible assignments and in providing students with technical training and support (notably during peak times near deadlines). To meet the demands of more advanced courses and students, the Center devotes considerable resources to training its own staff in a wide range of software applications and so can offer consultations and even informal tutorials on more specialized topics. Additionally, staff learn to assist students in research design and in the interpretation and presentation of their results. With this infrastructure in place, economics faculty teaching upper-level electives now assign their students more ambitious empirical research projects, and our thesis students are inclined to pursue more ambitious empirical topics.

## WORKING WITH THE WRITING PROGRAM DIRECTOR TO ADVANCE QUANTITATIVE REASONING

At Carleton College, two broader general education objectives are brought together through a partnership between the Economics department and the Writing Program: quantitative

reasoning (QR) and writing across the curriculum (WAC). For the last decade Carleton's Quantitative Inquiry, Reasoning, and Knowledge (QuIRK) initiative has supported faculty development to enhance students' capacity to create and evaluate arguments involving quantitative evidence. Working with the Writing Program enables Economics faculty to design writing assignments that teach and enforce QR skills, while tapping into the Writing Program's wealth of knowledge concerning best practice in assignment design.

Steen (2001, p. 1) brings the educational problem into sharp focus: "The world of the twenty-first century is a world awash in numbers….Unfortunately, despite years of study and life experience in an environment immersed in data, many educated adults remain functionally innumerate." DeLong and Rasky (2006, p. B15) argue that economists share in the responsibility of addressing this gap. To exemplify the problem, they critique a late 2005 *New York Times* story about a federal budget compromise. The newspaper explained that the Senate had passed "a sweeping five-year plan" to "reduce the deficit and save roughly $35 billion." The authors (and DeLong in particular) question whether such a bill can reasonably be described using words like "sweeping" or, as the article does later, "ambitious." From DeLong's perspective, "$35-billion over five years is 0.3 percent of federal spending, or 0.06 percent of gross domestic product. The spending cuts save $27 per person per year in the context of the federal government's current spending of $9,000 per person per year."

Note two characteristics of this example. First, it illustrates some of what economics students can offer to societal debates. Because we routinely contemplate a particular set of topics—the size of the US economy, tax levels and rates, government spending, median incomes, measures of income inequality—we can provide unique context to many questions. Second, note that the mathematics on display in DeLong's analysis is quite basic. He makes sharp use of basic

counts (the size of federal spending) and ratios (the deficit reduction relative to the US budget and relative to the US population). He also makes several shrewd rhetorical decisions when he chooses to frame the $35 billion reduction on a per capita basis. In a world where many readers confuse millions with billions or even trillions, such rhetorical choices are as critical as mathematics to the effectiveness of DeLong's argument.[3] In other words, as important as Calculus and marginal analysis are to the economists' habit of mind, our contribution to the national debate often rests as much on our ability to find, manipulate, and present quantitative data using basic mathematics—or in the words of Steen (2004, 9), "sophisticated reasoning with elementary mathematics more … than elementary reasoning with sophisticated mathematics."

While the problems of innumeracy are easy to see, the decision to address them with deliberate instruction comes with significant costs. A complete treatment would introduce students to commonly referenced data sources like the Economic Report of the President, the St. Louis Fed's FRED database, and the Statistical Abstract of the United States. Moreover, having armed students with data we would ask them to critically assess claims made by others and construct arguments of their own supported by quantitative evidence in text, table, and chart. And for students to learn to communicate arguments with quantitative evidence they will need opportunities to actually do so in more than a few classes. It is hard to do all of that without writing assignments, which raise obvious workload questions for students and faculty alike.

The good news is that we economists are not the first to encounter the workload issues surrounding writing assignments. Writing Program Directors have deep experience (often with other departments) developing strategies to manage the problem. (See Bean 2011, for example.) We have written elsewhere about the political advantages for a QR initiative to join forces with an existing writing program (Grawe and Rutz 2009). Here we share some ideas that one

economist picked up from a writing professional to improve his introductory microeconomics course.

**Example: Principles of Microeconomics Class**

The first lesson learned is that small can be powerful. Even though we aim to improve students' abilities to write persuasive arguments, in larger sections it may be impossible to assign many (or any) writing assignments. Short, ungraded assignments or brief class discussions on topics like how to read a figure or a table can lay the foundation for better work in later assignments or classes. Because students take more than one course, it is better to dispose of the "all or nothing" idea that we either teach all of writing or none at all. Thinking on the margin really is preferable!

Of course, if students are to learn to craft arguments with numbers we do have to assign major paper assignments in some contexts. We have done this in principles of microeconomics through a pair of papers about immigration policy. The first assignment asks students to use data from the Statistical Abstract to write a two-page op-ed arguing that immigration has deep effects in our economy such that whatever we do or don't do by way of reform is important. The second assignment requires them to write a four-page policy brief for a US Senator considering an immigration reform bill. The brief must include a supply-demand analysis that is contextualized by data in text, table, and graph.

Behind this assignment pairing lie two important lessons learned from the Writing Director. First, writing is a process. While principles courses are often too large to allow for students to submit drafts of papers for faculty feedback, by "scaffolding" assignments—having the lessons learned in the first re-occur in the second—students experience the opportunity to

apply lessons from the first paper to the second. While not "revision" exactly, by calling students' attention to this design feature the assignment scaffold allows similar opportunities for student growth. This choice has a second, practical benefit for grading. When students first write with quantitative evidence they predictably make mistakes. In one example, a student claimed that the number of unemployed in the early 2000s sat at roughly 7,000 because he failed to see that the table presented numbers measured in 1,000s. In another, a student mistook a table of "government consumption" for government expenditure (including transfers). No one said learning was easy! If the student's first encounter with manipulation and communication of data is combined with the first effort to present a supply-demand analysis, cognitive overload can result. Some students are simply overwhelmed by so many new tasks taken on at one time. Even previously mastered skills like paper organization suffer. However, student frustration is nothing compared to that of the faculty member who must sort out how to grade a nice supply-demand analysis embedded in embarrassing data work as compared with a weak theory explication coupled with exemplary data-rich context. Scaffolding the assignment avoids many of these challenges.

The second principle in the design of the assignments is even more critical: these papers helped the instructor to better achieve the learning goals for the course. Students grow in their ability to analyze policy with an economic model and to situate that analysis in real world context in order to make informed decisions. And that, after all, is at the heart of what most of us want to achieve in principles classes which mark the end of economics education for a majority of students. To be sure, teaching this assignment presents costs. Instructing students in how to find, manipulate, and present quantitative evidence takes approximately one-half class period. Even with the aid of grading rubrics, assessing student work and writing formative

comments on each paper adds to the term's grading load. Because the end of the term always comes before every economic topic is explored in all its glory, assignments must meet a high bar to survive from one term to the next. The reason this assignment keeps finding its way onto the syllabus is because it is authentic to course goals—it promotes academic growth so that students better achieve the course's goals.

In a number of economics courses we study how skill-biased technological change has altered economic returns and incentives. One manifestation of this theory can be seen in the dramatic reduction in the costs of finding and manipulating data. These skills are native to our discipline and so we have a distinct opportunity to empower all of our students—majors and not—to engage the data-drenched 21st century with skill and confidence. At Carleton, partnership with writing professionals has been a key part of bringing these aspirations to the classroom.

## THE LIBRARY DATA LAB AT THE UNIVERSITY OF CALIFORNIA, BERKELEY

Conducting research, often for a capstone course, is an important aspect of an undergraduate economics student's experience (Li and Simonson 2016, McGoldrick 2008a, McGoldrick 2008b). Students preparing empirically-based research papers often turn to their faculty for help with downloading and cleaning data sets. But very few faculty members are familiar with dozens of data sets. If students are limited to data sets in the faculty member's knowledge set, their paper topics may also be limited. Students will learn research skills but their projects might not follow their passion.

Into this world, enter the data librarian. A data librarian is someone with a background in both programming and social science research, who is available for student consultation and

course-related instruction. Data librarian positions in academic libraries are an outgrowth of the 1970s "Government Documents Librarian." Growth in the number of data librarians accelerated in the 1990s when libraries began to collect an increasing amount of material in digital formats. Libraries needed staff with the right skills to process, manage, and assist users with these digital materials. Today, data librarians are found at colleges and universities of all sizes and ranks, in the U.S., Europe, and beyond.

At UC Berkeley, the data librarian provides assistance that is split fairly evenly between more traditional research assistance offered to students on one hand and, on the other hand, identifying and supporting access to available datasets and providing technical assistance with software packages like R and Stata.[4] The types of assistance provided include file format conversions, data restructuring and merging operations, and technical documentation interpretation.

The data librarian works out of the Library Data Lab which provides individual and group study space for up to 20 students and houses several workstations that contain all the major statistical and GIS software packages. The Lab's unique niche is to support advanced undergraduates and to support instruction in courses requiring an empirical research project.

Without a data librarian, many students will start their data search with Google. The challenge is more than knowing the right key words to search. Data are available from multiple sources, in aggregated or microdata format, as individual downloadable files or in database systems. Moreover, students often encounter file formats that they are unfamiliar with (.dta, .sps, .dbx) and that don't cooperate when "double-clicked." Encountering "pay walls" or "privilege walls" is common. One solution is for the faculty member to provide everyone with a link to a specific dataset. With a data librarian, the options are much greater.

The data access landscape is constantly changing. Even faculty familiar with data sets used in their own research may find that their skills accessing unfamiliar data quickly atrophy. Curated sites like ICPSR, IPUMS, FRED, and various government agency sites evolve. Many libraries are increasing their spending on specific datasets sold by vendors or on subscriptions to databases such as Datastream (economic/financial), Data Planet (U.S. aggregate data), Roper (public opinion), or CEIC Global (international aggregate data). Emerging sites like Dataverse provide virtual collection datasets published by individual researchers or institutions. Many researchers continue to upload their data to personal or departmental websites. With such a vast array of constantly changing data sources, keeping up with change is far easier for librarians than faculty.

Comparative advantage is a powerful incentive. Faculty can focus on content, theory, and methodology while the data librarian focuses on downloading, cleaning, and merging data. Moreover having a physical on-campus space – a Lab – is a further advantage; students can do their own data work in proximity to a data librarian. And with the assistance of a knowledgeable data librarian, students are more likely to find data that is pertinent to their interests. Student passion for the topic goes a long ways toward creating an interesting, well-executed, and engaging research paper.

One concern that faculty sometimes express is a bleeding of data prep help into analysis help. Most librarians know when to draw the line and point the student back to their faculty member or graduate TA. Although some institutions employ highly trained librarians, perhaps with a PhD in a quantitative social science or strong programming backgrounds who can provide help beyond the basics, Berkeley's Library Data Lab does not provide assistance with statistical methodology. Questions about the model, the econometric approach, clustering standard errors,

interpreting coefficients, and so on are referred back to the faculty. This approach seems to work well, particularly since students who have received assistance with data preparation can return to their professors with "analysis-ready" data. Faculty office hours can focus on theory and methods, rather than data downloading and manipulation.

### Examples: Economics of Discrimination Research Seminar

The data librarian at Berkeley has provided assistance to Olney's upper-division research seminar in the Economics of Discrimination for several years. Students write a 15-20 page term paper on a topic of their choice which focuses on an issue of difference based on race, gender, sexual orientation, nativity, or other delineation. They must download, clean, prepare, and analyze relevant data. Students enter the course having completed Econometrics and knowing some basics of Stata or R.

Before the library hired a data librarian, Olney spent hours each week with students when they were at the data-downloading stage. Data sets on ICPSR, downloadable through DataFerrett, or directly off a government website were mostly fair game. But any topic that strayed beyond Olney's knowledge set had to be ruled out. The papers were fine but not terribly exciting. The students learned research skills adequately. Very few students were particularly excited about their paper.

With the data librarian, the paper topics are much richer because the data sources are better. One student, David Weinzimmer (2010), asked whether the extent of difference in wages between straight and gay workers varied with the degree of anti-gay sentiment in the worker's community, as proxied by the extent of the "yes" vote on California's 2008 Proposition 8 which banned gay marriage. The data librarian helped David merge two data sets: county-level election

results from the Statewide Database and wage and workplace data from the American Community Survey Public Use Microdata Sample. Creating the dataset involved mapping Public Use Microdata Areas (PUMAs) to counties. Predefined tables to map the two geographies to each other did not exist, so they had to do various merges with Census documentation and Statewide Database geographies. There is simply no way this project would have unfolded as it did without a data librarian.

Another student, Samantha Strimling (2014), merged three data sets to examine the extent of subprime lending by race: HUD's list of subprime lenders, HMDA data (using the unique respondent ID corresponding to each lender in the HMDA dataset), and the All-Transactions Housing Price Index, collected by the Federal Housing Finance Agency (FHFA). She then merged this multifaceted dataset with tract-level data from the 2000 U.S. Census in order to provide information on how each particular geographic area was perceived by lenders, which may (or may not) be different than the demographic composition of borrowers within a particular tract. As Samantha described the process at the time, she "basically lived in Harrison's lab for a few weeks."

Even when a project does not involve merging data sets, the data librarian's knowledge and guidance can make a big difference. Alexandra Fahey's paper (2014) on "Driving While Black" asked whether the incidence of drivers being stopped, being cited if stopped, and being searched if stopped varied systematically with the races of the driver and the officer. She used the Police-Public Contact Survey, a supplement to the National Crime Victimization Survey (NCVS). Without the data librarian, it is far less likely she would have found that data set.

Students frequently make creative use of Census (and American Community Survey) microdata. On several occasions, students have chosen topics involving the comparison of

demographic characteristics of married and unmarried partners. This particular type of project exemplifies the value of the services provided by the data librarian. Technically, in terms of the commands required, creating this data set is not particularly difficult. What is required, though, is a detailed familiarity with how to use a codebook to identify the variables needed to filter and match records, and a good grasp of the importance of variable naming. These skills are not typically covered in the curriculum, but are common when working with real world data.

Librarians are trained in the art of what the profession calls the "reference interview." Such a conversation can involve teasing out meaning from partially formed questions, assessing a student's level of knowledge and technical expertise and providing research advice accordingly, recognizing when a student's enthusiasm is tending towards biting off too much, and nudging them in a different direction when their topic is impossible or beyond their expertise. Librarians are trained to provide ideology-neutral and non-judgmental support. Students are not graded by the librarian and thus there is often less intimidation and more comfort asking what a student may fear is "a dumb question."

## Hiring a Data Librarian

Perhaps we've convinced you of the value a data librarian would add to your institution. What is your next step? First, familiarize yourself with the professional association for data librarians, iASSIST (International Association for Social Science Information Services &Technology, http://www.iassistdata.org). Next, form alliances across the social sciences in your institution. Libraries still tend to focus more on the humanities, and librarians predominantly have humanities backgrounds and interests. But a data librarian can benefit students and faculty across the social sciences.

Then, approach your Head Librarian. There is a definite trend of libraries creating data librarian positions, but not all of these positions provide the types of services described in this paper. Some focus more on helping researchers publish their data. Ideally, you can help the library craft an appropriate job description for your institution, but it should be assumed that some of the requirements are likely to be unfamiliar to a traditional library administrator. For this reason, a valuable resource at this stage is the job posting repository on the iASSIST website. For instance, a recent job listing from Vassar College stated "the librarian serves as a resource for working with data sources common to Social Sciences research, such as IPUMS, ICPSR, and other statistical databases ... [S]uccessful candidates will be well-versed in creating, interpreting, and manipulating statistical data using tools such as GIS, SPSS, Stata, and/or R." The resources and skills listed here are likely to be unfamiliar to a traditional librarian.

Keep in mind that there is a fairly small pool of qualified candidates on the market, particularly when a Masters of Library Science (MLS) degree is a requirement. One strategy is to write the position in a way that might appeal to candidates interested in an alternative academic ("alt-ac") career path. A data librarian job offers a path to remaining in academia with academic rank and, at some institutions, with faculty status. Positions with academic rank usually require an additional MLS degree but some employers will pay for this education. MLS degrees earned part-time in online programs are not necessarily valued less than traditional degrees, particularly for those with a relevant advanced degree or technology skills.

One solution sometimes called the "accidental data librarian" approach has worked for some institutions (Brown, Wolski, Richardson 2015). An existing librarian, ideally one with some academic preparation or other aptitude for the role can be offered the opportunity to retrain. For new data librarians or experienced ones who are updating their job skills, ICPSR annually

offers a highly regarded, weeklong, summer program in Ann Arbor, Michigan. The workshops offered at the annual iASSIST conference are another popular venue for picking up the necessary data skills.

## CONCLUSION

Student learning is enhanced when faculty partner with other campus professionals. In this article we have described three such partnerships: the Empirical Reasoning Lab at Barnard College, the College Writing Program and its role in the Quantitative Inquiry, Reasoning, and Knowledge initiative at Carleton College, and the Library's Data Lab at U.C. Berkeley. We encourage our colleagues to seek out similar existing partnerships at their own institutions, and to work with administrators to establish equivalent programs if they do not yet exist.

## NOTES

[1] We distinguish between empirical and quantitative reasoning, because in the context of general education requirements the latter term can often refer to formal analytical (e.g., mathematical) analysis. For a provocative view on the meaning of quantitative reasoning, see Neil Lutsky (2014), "Cut to the QUIC: What is the Essence of Quantitative Reasoning and How Can We Assess It?" (Unpublished manuscript, Carleton College). Empirical reasoning is similar to the notion of "Quantitative Literacy," found for example in the AAC&U (2010) VALUE -- Valid Assessment of Learning in Undergraduate Education – rubrics.

[2] For information on Barnard's T&L Center, see http://barnard.edu/about/teaching-learning-center. The ERC will join two other innovative learning spaces to constitute a "digital commons."

[3] For just one example of the challenge we face in this regard, the online version of a January 26, 2012 *New York Times* article includes a correction for misstating the probability of a virus mutation as "roughly one in a thousand billion billion" instead of "one in a thousand trillion trillion." Thanks to Mija Van Der Wege for sharing this example with us.

[4] Akers and Doty (2013) describe the role of the data librarians at Emory University, noting the key differences by discipline in faculty expectations and understandings of research data management needs.

## REFERENCES


Association of American Colleges and Universities (AAC&U). 2010. Quantitative Literacy VALUE Rubric. https://www.aacu.org/value/rubrics/quantitative-literacy (accessed July 15, 2016).

Akers, Katherine G. and Jennifer Doty. 2013. Disciplinary Differences in Faculty Research Data Management Practices and Perspectives. *International Journal of Digital Curation* 8 (2): 5-26.

Anselin, Luc. 1995. Local Indicators Of Spatial Association-LISA. *Geographical analysis*, 27 (2): 93-115.

Bean, John C. 2011. *Engaging Ideas: The Professor's Guide to Integrating Writing, Critical Thinking, and Active Learning in the Classroom*. 2nd ed. San Francisco, CA: Jossey-Bass.

Brown, Rebecca A., Malcolm Wolski, and Joanna Richardson. 2015. Developing New Skills for Research Support Librarians. *Australian Library Journal* 64 (3): 224-234.

Delong, J. Bradford and Susan Rasky. 2006. "He Said, She Said," *Chronicle of Higher Education* 52 (33): B15.

Fahey, Alexandra. 2014. "Driving While Black in the United States: Do Police Officers Use Stop and Search as a Discrimination Tactic?" Unpublished manuscript. Economics 154 Term Paper, University of California, Berkeley.

Grawe, Nathan D. and Carol A. Rutz. 2009. Integration with Writing Programs: A Strategy for Quantitative Reasoning Program Development. *Numeracy* 2 (2): Article 2.

Hensley, Merinda Kaye, Sarah L. Shreeves, and Stephanie Davis-Kahl. 2014. A Survey of Library Support for Formal Undergraduate Research Programs. *College & Research Libraries* 75 (4): 422-441.

Li, Ishuan and Robert Simonson. 2016. Capstone Senior Research Course In Economics. *The Journal of Economic Education* 47 (2): 161-167.

Lutsky, Neil. 2014. Cut to the QUIC: What is the Essence of Quantitative Reasoning and How Can We Assess It? Unpublished manuscript. Carleton College. https://apps.carleton.edu/quirk/assets/aalac_lutsky_slides.pdf (accessed April 5, 2017).

McGoldrick, KimMarie. 2008a. Doing Economics: Enhancing Skills Through A Process-Oriented Senior Research Course. *Journal of Economic Education* 39 (4): 342–56.

McGoldrick, KimMarie. 2008b. Writing Requirements And Economic Research Opportunities In The Undergraduate Curriculum: Results From A Survey Of Departmental Practices. *Journal of Economic Education* 39 (3): 287–96.

Pace, Robert K., and LeSage, James P. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman and Hall/CRC.

Pear, Robert. 2005. Senate Passes Budget with Benefit Cuts and Oil Drilling. *New York Times*, November 4, 2005. https://nyti.ms/2oK8CfS (accessed April 5, 2017).

Steen, Lynn Arthur. 2001. *Mathematics and Democracy: The Case for Quantitative Literacy*. Washington, DC: Mathematical Association of America.

Steen, Lynn. 2004. *Achieving Quantitative Literacy*. Washington, DC: Mathematical Association of America.

Strimling, Samantha. 2014. "Race, Ethnicity, and Subprime Lending: A National Level Analysis." Unpublished manuscript. Economics 154 Term Paper, University of California, Berkeley.

Weinzimmer, David M. 2010. "Is There Evidence of Wage Discrimination Against Gays and Lesbians?" Unpublished manuscript. Economics 154 Term Paper, University of California, Berkeley.

Zimmer, Carl. 2012. Study Finds Virus to Be Fast Learner on Infecting. *New York Times*, January 26, 2012. https://nyti.ms/2oKaXHL (accessed on August 25, 2015).